

Our Docket No.: 51876P380
Express Mail No.: EV339917661US

UTILITY APPLICATION FOR UNITED STATES PATENT
FOR
APPARATUS AND METHOD FOR DYNAMIC RESOURCE ALLOCATION IN
INTERACTIVE SATELLITE MULTIMEDIA SYSTEM

Inventor(s):
Ki-Dong Lee
Yong-Hoon Cho
Ho-Jin Lee

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 Wilshire Boulevard, Seventh Floor
Los Angeles, California 90025
Telephone: (310) 207-3800

APPARATUS AND METHOD FOR DYNAMIC RESOURCE ALLOCATION IN
INTERACTIVE SATELLITE MULTIMEDIA SYSTEM

Field of the Invention

5

The present invention relates to an apparatus and method for resource allocation in interactive satellite multimedia system; and, more particularly, to a dynamic resource allocation apparatus and method for assigning
10 timeslot in a return channel in multi-frequency time division multiple access (MF-TDMA) in order to have a maximum throughput.

Description of Related Arts

15

There have been various apparatus and methods developed for resource allocation in a time division multiple accesses (TDMA) such as a timeslot assignment method in inter-satellite links and a resource allocation
20 method in TDMA in mobile communication system.

Inhere, a resource represents a frequency bandwidth used at a return link, which is a link between a return channel satellite terminals (RCSTs) to a satellite. That is, the resource is a set of timeslots in the return link
25 in multi-frequency time division multiple access MF-TDMA. In MF-TDMA, resources are allocated as a unit of a frame or a super-frame. One frame includes a plurality of timeslots

and one super-frame includes a plurality of frames.

An interactive satellite multimedia system is a satellite network with one earth station (HUB) and a plurality of terminals called return channel satellite terminals (RCST). The earth station (HUB) receives a request of service through satellites and provides services to terminals (RCST) through the satellite by responding to the request. In the satellite communication system, since the radio resources are very expensive comparing to those in land mobile communication systems, it is one of the most important problems to maximize the link throughput using the limited available resources.

Specially, in the interactive satellite communication system having a centralized resource management function in the earth station (Hub), the resource needs to be allocated by considering round-trip time (RTT) spending for requesting the resource from the terminal (RCST) to the earth station (Hub) and receiving an assignment plane from the earth station (Hub) to terminal station (RCST). In a view of allocating resource by considering RTT, the resource allocation method in the interactive satellite communication system is distinguished from the conventional technique for allocating resources in the satellite communication.

Furthermore, comparing to a timeslot assignment scheduling and packet assignment scheduling of a mobile communication system, it is also distinguished in view of

delay of transmission and a bandwidth of frequency is comparatively broadband. In broadband TDMA, the length of each timeslot is usually much shorter than that in a narrow- or medium- band TDMA. As a result, the scheduling period is getting short or the number of timeslots per a scheduling period is getting great, both of which make a fast timeslot scheduling more difficult, especially in such an interactive satellite network having an intrinsic round trip time between capacity request and allocation.

As a result, it is preferred to have as short period of scheduling time as possible in order to minimize the negative effect of the RTT.

In the interactive satellite multimedia system, therefore, it is essential that a method for reapidly and dynamically outputting a resource allocation plan to capacity request from a plurality of terminals (RCSTs).

Summary of the Invention

It is, therefore, an object of the present invention to provide an apparatus and method for rapidly generating a dynamic resource allocation schedule in order to maximize throughput in a return link based on MF-TDMA of an interactive multimedia network.

In accordance with an aspect of the present invention, there is provided an apparatus for dynamically allocating resource in an interactive satellite multimedia system,

including: resource request amount collection unit for
accumulating a requested resource amount corresponding to
each of terminals during a super-frame period; resource
request amount processing unit for dividing an accumulated
5 requested resource amount by the number of frame pairs in a
super frame and storing a sum of a result of dividing and
rounding up a remain of the division to a nearest integer
as a request amount of each corresponding terminal ; and
resource allocation unit for deciding a time slot allocated
10 at each of terminals corresponding to a frame pair based on
optimal allocation amount, which is decided based on the
request amount by the requested amount processing unit.

In accordance with an aspect of the present invention,
there is also provided a method for dynamically allocating
15 resources in an interactive satellite multimedia system,
including the steps of: a) accumulating a request amount of
resource corresponding to each of terminals during a super-
frame period; b) dividing the accumulated request amount of
resource by the number of frame pairs included in one super
20 frame and remembering a sum of a result of dividing and
rounding off a remain of the division to a nearest integer
as a resource request amount; and c) deciding an optimal
allocation amount based on the resource request amount and
deciding a time slot to be allocated to each of terminals
25 based on the optimal allocation amount.

In accordance with an aspect of the present invention,
there is also provided a computer readable recording medium

storing instructions for executing a method for actively allocation resource in two-way satellite multimedia system, including functions of: a) accumulating a request amount of resource corresponding to each of terminals during a super-
5 frame period; b) dividing the accumulated request amount of resource by the number of frame pairs included in one super frame and remembering a sum of a result of dividing and rounding off a remain of the division to a nearest integer as a resource request amount; and c) deciding an optimal
10 allocation amount based on the resource request amount and deciding a time slot to be allocated to each of terminals based on the optimal allocation amount.

Brief Description of the Drawing(s)

15

The above and other objects and features of the present invention will become apparent from the following description of the preferred embodiments given in conjunction with the accompanying drawings, in which:

20 Fig. 1 is a diagram showing a conventional interactive satellite multimedia system;

Fig. 2 is a block diagram depicting a dynamic resource allocation apparatus in accordance with a preferred embodiment of the present invention;

25 Fig. 3 is a flowchart for explaining a dynamic resource allocation method in accordance with a preferred embodiment of the present invention;

Fig. 4 is a diagram showing a structure of a super-frame in accordance with a preferred embodiment of the present invention; and

Figs. 5A to 5D show results of extensive simulation of a preferred embodiment of the present invention.

Detailed Description of the Invention

Other objects and aspects of the invention will become apparent from the following description of the embodiments with reference to the accompanying drawings, which is set forth hereinafter.

Fig. 1 is a view representing a conventional interactive satellite multimedia system.

Referring to Fig. 1, the conventional interactive satellite multimedia system includes a plurality of terminals 11, a satellite 12 and a host station (HUB) 13. The terminals 11 send a request for resource to the host station 13 through the satellite 12. The host station 13 provides services to the terminals 13 through the satellite 12. In the interactive satellite multimedia system, it is one of the most important problems to maximize a throughput according to the user's traffic.

In Fig. 1, there is only one host station and one satellite however, it is possible to include multiple host stations and satellites in the satellite communication network.

The multiple terminals 11 simultaneously request resources to the host station 13 and each of terminals uses only allocated resource (timeslot).

5 The host station 13 receives and analyzes requests for resources from the multiple terminals and decides appropriate number of timeslots to assign to each of terminals according to analysis. The decision of the host station is reported to each of terminals through the satellite 12.

10 The host station includes an algorithm for computing a dynamic resource allocation schedule in order to maximize the throughput of a return link between the terminal 11 to the satellite 12. The host station receives requests for resource of frequency from the terminals 11 and rapidly
15 computes the dynamic resource allocation schedule. At this time, a requested resource for a frequency is a frequency bandwidth used for a return link between the terminal 11 and the satellite 12. That is, the requested resource for frequency is a set of a certain number of timeslots in the
20 return link based on MF-TDMA. For rapidly computing a schedule for allocation resources at each of superframes, the present invention needs to improve the computational efficiency of timeslot scheduling by employing the so-called problem reduction and decomposition technique, where
25 we separately consider how to decide an amount of allocation according to each of terminals and how to select timeslot for allocating to each of terminals.

By using a problem decomposition method, the amount of computation can be decreased and the speed of computation can be improved.

Fig. 2 is a diagram showing a dynamic resource
5 allocator for dynamically allocating resources in accordance with a preferred embodiment of the present invention.

Referring to Fig. 2, the dynamic resource allocator in the host station (Hub) includes a resource request
10 collector 21, a resource request amount processor 22 for computing an amount of resources for allocation by dividing collected request amount to the number of frame pairs and rounding off a result of division to nearest integer number, a resource allocation amount decider 24, a resource
15 allocation scheduler 25 and a resource allocation schedule copier 26.

That is, for scheduling the resource of the first frame pair, the resource allocation decider 24 decides the optimal amount for each frame pair by considering how many
20 frames are allocated to each of terminals 11. The resource allocation scheduler 25 makes a schedule based on the optimal amount of allocation and selects one of terminal to be assigned to allocated resource. The resource allocation schedule copier 26 copies allocation schedule of the first
25 frame pair to remained frame pairs in the same superframe.

Referring to Fig. 2, the dynamic resource allocator used for the host station 13 includes a resource request

collector 21 for accumulating an amount of request for resource during a superframe period, a resource request amount processor 22 for computing a requested resource amount by dividing collected request amount to the number of frame pairs in one of superframe and adding the result of dividing and rounding off a remain of division to nearest integer number and a resource allocator 23 for deciding an optimal amount of allocation based on the request amount of the request amount processor 22 and deciding a timeslot to be assigned to each of terminals 11 according to the frame pairs based on the optimal amount of allocation.

The resource allocator 23 has a resource allocation amount decider 24 for deciding the optimal amount based on processed request amount in the resource request processor 22, a resource allocation scheduler 25 for selecting and deciding one of terminals RCST 11 to be allocated of a timeslot included in the first frame pair based on the optimal allocation amount and a resource allocation scheduler copier 26 for copying timeslot allocation schedule of the first frame pair to remained frame pairs in the same superframe.

The resource request collector 21 accumulates a requested resource amount generated during the superframe according to each of terminals RCST and reports it to the resource request amount processor 22. After reporting, the accumulated amount is set to 0 and the resource request

collector 21 starts to accumulate the requested resource amount with next superframe.

The resource request amount processor 22 computes an optimal amount of allocation by dividing a resource request
5 amount of each of terminals RCST 11 to the number of frame pairs and rounding off a result of division to the nearest integer number and the resource allocation decider 24 decides the optimal allocation amount according to priorities such as cost and weight.

10 The resource allocation scheduler 25 decides terminals according to each timeslot included in the first frame pair based on the optimized allocation amount.

The resource allocation schedule copier 26 copies schedule of resource allocation including a timeslot
15 allocation schedule from the resource allocation scheduler 25 assigned to the first frame pair to remained frame pairs in same super-frame.

Fig. 3 is a flowchart explaining dynamic resource allocation method in accordance with a preferred embodiment
20 of the present invention.

At first, a unit of resource allocation according to a schedule is defined at step of 301. That is, a super-frame and frame pair are defined.

In MF-TDMA, the schedule for allocation is
25 established based on a unit of frame in the superframe and one super-frame includes a plurality of frames.

A frame pair is a subset of the superframe and it is

satisfied by below conditions.

A frame pair is a set of frame having identical time period. Each of frame pairs has same arranging type of time slot and a frequency bandwidth and only difference between frame pairs is the time period. Therefore, all frame pair is a subset of the super-frame and a sum of all frame pairs defined in a given superframe constitutes a superframe.

For example, in case of the superframe having 32 frames as shown in Fig. 4, there are 16 frame pairs by pairing frame 0 and frame 16 as a one frame pair and so on.

After defining the frame pair or superframe, the resource request amount collector 21 remembers an ID of terminals, which are logged on at step of 302, and accumulated request amount during one super-frame length is stored at step of 303. At the step of 303, the request amount is updated. The accumulated request amount of each terminal is computed by dividing the accumulated request amount to the number of frame pairs in one super-frame and adding the result of division and rounding up the remain of the division to nearest integer number. A small difference will be ignored since each request amount is reputedly allocated as many as the number of frame pairs.

Next, the resource allocation decider 23 computes the optimal allocation amount based on request amount at step of 304 and the resource allocation scheduler 25. The resource allocation copier 26 decides timeslots for

allocation to each terminal according to frame pair based on the optimal allocation amount at step of 305 and 306.

At step of 305, the resource allocation scheduler 25 decides all time slots including the first frame pair to
5 assign to which a terminal (RCST) 11. At step of 306, a time slot allocation schedule of the first frame pair is copied to remained frame pairs in the identical super-frame by the resource allocation schedule copier 26.

For example, in case of the super-frame including 32
10 frames, as shown in Fig. 4, an algorithm of equations 4 and 5 is applied for deciding to allocate a time slot to a first frame pair by generating a resource allocation schedule and for other frame pairs, the timeslots are allocated by copying the resource allocation schedule for
15 the first frame pair. The above mentioned method for deciding the assignment of each time slot can reduce calculation steps and processing time. For example, if there are 2000 time slots in a frame, 2000×2 times of calculations for the first frame are required. That is,
20 there are 4000 allocation schedules required and the other frame pairs use the computed allocation schedule for the first frame. However, if allocation schedules for all frame need to be calculated then there are 64000 allocation schedules required.

25 Hereinafter, a step of 304 for deciding optimal allocation amount decision of the first frame pair in the identical superframe and a step of 305 for rapidly

allocating time slot based on the optimal allocation amount are explained in detail.

At first, it is explained that the optimal allocation amount based on an accumulated during a length of the superframe received from the terminals (RCST) 11, which is logged on.

Below equation shows original problem for dynamic allocation of time slots.

(CAP)

$$\text{MINIMIZE } g(x) = \sum_{j \in R} [v_i \sum_{m \in A} \sum_{i \in S} (1 - x_{ij})]$$

10 subject to

$$\sum_{m \in A} \sum_{j \in S} x_{ij} \leq \min\{Q_j, X_j + Y_j\}, j \in R$$

$$\sum_{m \in A} \sum_{j \in S} x_{ij} \geq Y_j, j \in R$$

$$\sum_{j \in R} x_{ji} \leq 1, j \in S_m, m \in A$$

$$x_{ji} \in \{0, 1\}$$

Eq. 1

In Eq. 1, i is i^{th} time slot, j is j^{th} terminal and R is a set of IDs of terminals logged on. A is a set of ID of return link demodulators, which are available. S_n is a set of IDs of available time slots and Q_j is the maximum number of time slots, which can be allocated to the terminal j and X_j is the number of time slots requested by terminal j , Y_j is the minimum number of time slots that should be allocated to terminal j . x_{ij} is a decision variable, which has a value of 1 when a time slot i is allocated to a terminal j and has a value of 0 otherwise. V_j is a penalty coefficient (penalty weight). The penalty

coefficient is decided according various factors such as delay of the residual packets on the transmission queue of each terminal (RCST). That is, a large value is assigned for the penalty coefficient to the terminal (RCST) when the terminal contains a lot of delayed data. As a result, fairness of allocation resources can be achieved.

Referring to Eq. 1, $g(x)$ is the total penalty caused by the timeslot allocation denoted by matrix $[x_{ij}]$. There are several conditions. At first, the number of time slot for allocating to each terminal must be less than a upper bound (threshold value) or request number, which is decided by system parameters such as construction of the super-frame. Secondly, the number of time slots for allocating to each terminal must be greater than lower bound, which is decided for satisfying service quality. Lastly, each time slot cannot be allocated to two and more terminals.

Eq. 1 can be modified to following Eq. 2 for effectively solving a problem of resource allocation based on a concept of frame pair in accordance with the present invention. The capacity allocation problem (CA)) can be mathematically formulated as follows:

(CAP)

$$\text{MINIMIZE } g_1(x) = \sum_{j \in R} v_j \sum_{m \in A} \sum_{i \in S} (1 - x_{ij})$$

subject to

$$\begin{aligned}
& \sum_{m \in A} \sum_{j \in S_{m,1}} x_{ij} \leq \min\{Q_j, X_j + Y_i\}, j \in R \\
& \sum_{m \in A} \sum_{j \in S_{m,1}} x_{ij} \geq Y_j, j \in R \\
& \sum_{j \in R} x_{ji} \leq 1, j \in S_m, m \in A \\
& x_{ji} \in \{0,1\}
\end{aligned}
\tag{Eq. 2}$$

A difference between Eq. 1 and Eq. 2 is that the number of objects for allocation schedule is a total number of slots in a frame pair and a total number of slots in a superframe, therefore, the processing time may be decreased since the size of problem (CAP) and the complexity required to solve problem (CAP) are decreased. Furthermore, if the allocation schedule of the first frame pair is copied and used for remained frame pairs, the allocation schedule for time slots in the super-frame can be effectively and rapidly made.

In Eq. 2, $S_{m,1}$ is a set of available time slots included in the first superframe. The original problem (CAP) is divided to n sub-problems and optimize x . As a result, $g(x) = \sum_{m=1}^n g_n(x)$ at the optimal solution x .

Eq. 2 requires a preprocess of solving an allocation decision amount.

A problem of deciding an amount of allocation v_j is a weight representing how a terminal j is important comparing to other terminal. The weight can be decided according to a service quality. For example, resources are effectively and rapidly allocated to the terminals by assigning higher

priority to a terminal, which have been delayed to be allocated of resources for long period time.

$$\text{MINIMIZE } f(x) = \sum_{j \in R} v_j z_j$$

Subject to

$$Z_j \leq \min\{Q_j, X_j + Y_j\}, j \in R$$

$$Z_j \geq Y_j, j \in R$$

Eq. 3

$$\sum_{j \in R} z_j \leq \sum_{m \in A} S_{m,1}$$

$$Z_j \in Z \cup \{0\}$$

5

Referring to Eq. 3, Z_j is a decision variable denoting the number of time slot allocated to the terminal j .

$f(z)$ denotes a total penalty according to the amount of allocation denoted by vector z . There are limitation conditions as followings. At first, an amount of allocation of each terminal must be less a given upper bound and request amount. Secondly, the amount of allocation of each terminal must be greater than a given lower bound and lastly, the total amount of allocation must be less than available amount of resource.

A problem for deciding an amount of allocation is shown in below.

Step 1 (Sort $[v_j]$)

$k := 1, J_{k-1} = \{\}$

For($k \leq |R|; k++$) $\{$

$j_k := \arg \max\{v_j, j \in R - J_{k-1}\}$

$J_k := J_{k-1} \cup \{j_k\}$

$\}$

Step 2 (Find an optimal y^*)

$n := \max\{n \mid \sum_{k=1}^n y_j k \leq \sum_{m \in A} |S_m| - \sum_{j \in R} Y_j\}$

if $n = |R|$, *then* $y_j := \min\{Q_j - Y_j, X_j\}$ *for* $j \in R$

Else

$y_j^* := \min\{Q_j - Y_j, X_j\}$ *for* $j \in J_n$

$y_m + 1 := \sum_{m \in A} |S_m| - \sum_{j \in R} Y_j$
 $\quad \quad \quad - \sum_{j \in J_n} \min\{Q_j - Y_j, X_j\}$

Step 3 (Find an Optimal Z^*)

Eq. 4

$Z^* := y^* + [Y_j]$

In a Step 1, V_j is sorted from one having large weight to one having small weight. In here, j_k is an ID of a terminal having k^{th} largest penalty weight (V_j). J_k is a set
5 having $\{j_1, \dots, j_k\}$.

In the Step 2, additional allocation amount is optimally decided. At $n := \max\{\dots\}$, the additional allocation amount is the number of variable timeslots $\sum x_{jk}$.
If an amount of remained resource is sufficient, additional
10 requested resource is maximally allocated and if the amount of remained resource is not sufficient, additional requested resource is maximally allocated to a terminal J_n having a highest priority at first and then remained resource is allocated to other terminals.

15 Step 3 is a step of deciding total allocation amount.

The total allocation amount is a sum of basic allocation amount and additional allocation amount and it is computed for each terminal.

Referring to Eq. 4, a step 304 for deciding optimal
5 resource allocation amount is explained in below.

At first, it is sorted from one having a higher weight and one having a lower weight. At this time, the highest weight has an index j_1 and stored in a set J_1 . Similarity, the k^{th} highest weight has an index j_k and
10 stored in a set J_k .

After then, an additional allocation amount y_j^* excepting minimum allocation amount is decided at step 2. At step 3, a total allocation amount z_j^* is calculated by adding the decided additional allocation amount and the
15 optimal allocation amount. In here, $y^*=[y_j^*]$, and $z^*=[z_j^*]$.

In other hand, below equation 5 shows a step 305 for deciding a time slot to be allocated to each of terminals RCST based on the optimal allocation vector z^* . That is, it is a step for choosing time slot to be allocated to each of
20 terminals based on the decided allocation amount $[z_j^*]$.

Step 1: Initialization

slot_count:=0, x:=0

Step 2(Iteration)

For (k:=1;k<=|R|;k++){

For (i:=slot_counter; i<+j_k;i++){

Eq. 5

xi,jk:=1;

}

slot_counter+=z_{jk}.

}

Referring to Eq. 5, a slot i is allocated to a terminal j_k, x_i and j_k are set as 1. That is, the i is increased as many as the number of allocation and x_i and j_k are set as 1, therefore, an amount of allocation to the terminal j_k is z_j^{*}.

In Eq. 5, objective time slots are all of time slots in a first frame pair. The time slot is allocated from one having lowest frequency (carrier) to a direction of time increase. The above mentioned step is repeated for next carrier.

A schedule of frame pair generated in steps of 304 and 305 is copied to remained frame pairs at step of 306. Therefore, allocation schedule for all time slots included in the super-frame is completed.

As mentioned above, a reason of computing the optimal allocation amount is since values of objective equation are not changed if it is moved to any directions.

Figs 5A to 5D shows a simulation result of the preferred embodiment of the present invention for verifying an effectiveness of the present invention. For running

simulation of the present invention, a personal computer with Pentium III 1.0 GHz is used. The simulation uses a model network having 128 terminals and 65024 time slots under conditions that all weights are 1, a minimum allocation amount is 6 per a frame, a maximum allocation amount is 508 per a frame. The simulation is progressed by randomly generating a request resource amount and measuring a time solving two problems of the request resource amount.

A program used for simulation has already used at a real-time resource manager in a host station Hub 13 and it includes a time for expressing a resource allocation schedule as a type of messaging between units of the real system. According to the result of the simulation, it maximally takes a time of 10 ms for deciding an allocation amount and also maximally takes a time of 30 ms for selecting a time slot. It also takes 500 ms for requesting resource of the terminal RCST 11 to a main station Hub 13, completing an allocation schedule and the allocation schedule is passed to the terminal RCST 11 again. Although a length of superframe is a range of 2000 ~ 92000ms (a standard of DVB-RCS), it is clearly shows that a processing speed of the present invention is very fast comparing to prior art.

The above-mentioned method according to the present invention can be implemented as instructions and stored to a computer readable recoding medium such as a CD-ROM, RAM, floppy disk, hard disk and optical magnetic disk.

As mentioned above, the present invention can output resource allocation schedule in order to have a maximum throughput in a timeslot allocation in a return link based on MF-TDMA of interactive multimedia network by a problem
5 division method using a frame pair.

While the present invention has been described with respect to certain preferred embodiments, it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the
10 scope of the invention as defined in the following claims.